



Prijsvraag

De prijsvraag ging de afgelopen keer over extreme waarden-theorie. Er waren twee kwalitatief goede inzendingen, waarbij één inzending uitblonk in grondigheid en uitgebreidheid. Niet minder dan 4 pagina's met afleidingen, toelichtingen en figuren werden door de winnaar ingezonden, die alle blijk gaven van veel inzicht en plezier bij het maken van de opdracht.

De prijsvraag is daarmee gewonnen door drs. Pieter Marres AAG. Hij is zelfstandig actuarial data scientist met zijn eigen bedrijf QuantSense. Zijn drijfveer is om organisaties betere besluiten te laten nemen met data en gestructureerde digitale processen. Daarnaast verzorgt hij met veel plezier trainingen in programmeren met Python. Wat hij interessant vond aan de prijsvraag is dat de vermeende heilige graal in de statistiek – namelijk dat een gemiddelde uiteindelijk altijd wel convergeert naar een normale verdeling – in de wereld van de extreme waarden niet op blijkt te gaan, zelfs wanneer je de steekproefomvang opvoert naar ongekende hoogte. Dit komt omdat de centrale-limietstelling een eindige variantie nodig heeft, waaraan niet wordt voldaan in de opgave.



Pieter Marres



John Einmahl

De prijsvraag is opgesteld door professor John Einmahl, emeritus hoogleraar statistiek aan Tilburg University. Hij geeft ook het onderdeel extreme-waardentheorie in de Executive Master of Actuarial Science. Naast zijn onderwijs kent iedere actuaaris hem vanwege de beroemde Dekkers-Einmahl-de Haan schatter voor de index van een extreme-waardenverdeling. Naast zijn fundamentele onderzoek staat hij ook bekend om zijn inspirerende toepassingen van extreme-waardentheorie voor een breed publiek. Aansprekende vragen die hij heeft beantwoord met extreme-waardentheorie zijn: hoe laag kan het wereldrecord voor 100 meter hardlopen worden? En extra relevant voor actuarissen: hoe oud kunnen mensen maximaal worden? Het antwoord op de eerste vraag luidt 9.36 seconden voor mannen en 10.33 voor vrouwen. Het antwoord op de tweede vraag luidt: niet meer dan ongeveer 125 jaar.

Ook in de gestelde prijsvraag wist John een verrassend voorbeeld te geven van de consequenties die extreme-waardentheorie met zich mee kan brengen.

UITWERKING

De prijsvraag ging over de dikstaartige Pareto-verdeling met de volgende verdelingsfunctie: $F(x) = 1 - 1/(1+x)^\beta$. De verwachtingswaarde van de Paretoverdeling wordt gegeven door: $\mu = E[X] = 1/(\beta - 1)$.

- a) We willen de waarde van c bepalen zodanig dat $P[X \geq c | \mu = 1000] = \alpha$ met $\alpha = 0.05$. Bij een observatie die deze threshold c overschrijdt, zal de test concluderen dat het zeer onwaarschijnlijk is dat de verwachtingswaarde (maximaal) gelijk is aan μ . Op basis van de verwachtingswaarde μ onder de nulhypothese leiden we af dat $\beta = 1.001$. We willen daarom oplossen dat $1 - F(c) = \alpha$, oftewel: $1/(1+c)^\beta = \alpha$. De oplossing wordt gevonden door $c = 18.94$. We concluderen dat een observatie hoger dan 18.94 al tot een verwerping leidt van de hypothese dat de verwachtingswaarde (maximaal) gelijk is aan 1000.
- b) Vervolgens berekenen we de kans $P[X \geq c | \mu = 2000]$ dat de ontwikkelde test wordt verworpen wanneer de werkelijke verwachtingswaarde wordt gegeven door $\mu = 2000$. Voor deze alternatieve verwachtingswaarde leiden we af dat $\beta = 1.0005$. We vinden dan dat $1 - F(c) = 1/(1+c)^\beta = 1/(1+18.94)^{1.0005} = 0.05$. De power van de statistische test is vrijwel gelijk aan α , en daarmee slechts 5%. Oftewel, zelfs als de werkelijke verwachtingswaarde twee keer zo groot is als de nulhypothese veronderstelt, dan nog zal de nulhypothese waarschijnlijk niet worden verworpen.

Deze prijsvraag laat zien dat dikstaartige verdelingen tot contra-intuïtieve resultaten kunnen leiden voor statistische testen met betrekking tot de verwachtingswaarde.

De complete uitwerking staat op de website van het AG.